A Neural Network Approach for Truth Discovery in Social Sensing

Jermaine Marshall, Arturo Argueta, Dong Wang
Department of Computer Science and Engineering
University of Notre Dame
Notre Dame, IN 46556
jmarsha5@nd.edu, aargueta@nd.edu, dwang5@nd.edu

***Abstract*—Social sensing has emerged as a new application paradigm in networked sensing communities where a colossal amount of observations about the physical world are contributed by people or devices they use. Our work solves a critical challenge in social sensing applications where the goal is to estimate the reliability of social sensors and the truthfulness of observed variables (typically known as claims) with little prior knowledge on either of them. This challenge is referred to as *truth discovery*. An important limitation in the previous truth discovery solutions is that they assume the relationship between source reliability and claim truthfulness can be represented by *simplified functions (e.g., linear, quadratic and binomial)*. This assumption leads to suboptimal truth discovery results because the exact relational dependency between sources and claims *is often unknown a priori*. In this paper, we show that a neural network approach can learn the complex relational dependency better than the previous truth discovery methods. In particular, we develop a multi-layer neural network model that solves the truth discovery problem in social sensing *without any assumption on the prior knowledge of the source-claim relational dependency distribution*. The performance of our model is evaluated through two real-world events using data crawled from Twitter. The evaluation results show that our neural network approach significantly outperforms previous truth discovery methods.**

***Index Terms*—Neural Network, Deep Learning, Truth Discovery, Relational Dependency, Social Sensing, Twitter**

I. Introduction

In this paper, we develop a neural network approach to solve a pressing challenge in social sensing known as the truth discovery problem. Social sensing is a new application paradigm in networked sensing communities where observations of the physical world are gathered from humans and/or devices on their behalf (e.g., smartphones) [1]. Examples of social sensing applications include real-time disaster response using social media [2], intelligent transportation systems using data from mobile phones [3], and geotagging applications using inputs from common citizens [4]. Social sensing has clear advantages over the traditional sensing paradigms that use infrastructure sensors: (i) social sensing provides an inexpensive way to obtain data on a massive scale since it is infrastructure free; (ii) social sensors (e.g., humans) are versatile and able to observe a wide range of events (e.g., disasters, fires, riots, and traffic jams). However, a fundamental challenge for social sensing is to estimate the reliability of social sensors and the truthfulness of observed variables (typically known as claims) with little prior knowledge on either of them. This challenge is known as *truth discovery*.

Good progress has been made in developing methods for solving the truth discovery problem in information fusion [5], [6], machine learning [7], data mining [8], [9], and networked sensing areas [10], [11]. However, a common assumption made in previous works is that the relationship between the reliability of sources and claim truthfulness, which is referred to as *source-claim relational dependency*, can be represented by some simplified functions (e.g., linear [12], quadratic [13], binomial [10]). This assumption can easily lead to suboptimal truth discovery results because neither the claim correctness nor the source reliability is known *a priori*

In this paper, we develop a new truth discovery scheme based on *neural networks* to address the above limitation. Neural networks have been widely used in computer vision, image processing, and natural speech recognition to estimate or approximate complex functions that might depend on a large number of inputs. This nice feature of the neural network approach makes it especially suitable to accurately estimate the source-claim relational dependency function, which can be very complex. In particular, we develop a new truth discovery solution based on feed forward neural networks. In these models, brain neurons and synapses are represented by adjacency matrices, which are used to approximate functions that are not learned efficiently by regular models (e.g., logistic regression, decision tree, etc.).

However, a few challenges exist in developing a multi-layer neural network approach for solving the truth discovery problem in social sensing. First, it is challenging to decide the number of layers of the neural network model as the complexity and distribution of the source-claim relational dependency are often unknown a priori. Furthermore, it is not a trivial task to design an efficient learning algorithm to update the numerical parameters (e.g., adaptive weights of matrices) of the neural network model. Third, it can be difficult to decide what percentage of neurons should be kept or dropped out to avoid the over-fitting problem in the training process of the neural network model.

In this paper, we develop a multi-layer neural network model to accurately capture the complex source-claim relational dependency. In particular, a neural network for truth discovery is defined by a set of input neurons which are

978-1-5090-0898-8/16/$31.00 ©2017 IEEE

activated by the social sensing data (i.e., sources and the claims they make). After being weighted and transformed by a learning function, the activation of these neurons are then passed on to other neurons inside the neural networks. This process is repeated until the output neuron that determines the truthfulness of a claim is activated. The complex source-claim relational dependency is learned by the neural network model through the above training process. Our scheme is evaluated through two real-world events from the physical world using data crawled from Twitter: Brussels Bombing event (2016) and Baltimore Riots event (2015). The evaluation results show that our neural network approach significantly outperforms previous truth discovery methods in terms of both effectiveness and efficiency in social sensing applications.

Our contributions are summarized below:

- To the best of our knowledge, we are among the first to develop a *neural network method* for solving the truth discovery problem in social sensing.
- The proposed neural network approach can accurately capture the source-claim relational dependency *without casting it to some over-simplified functions*.
- We show strong performance gains in our evaluation achieved by the neural network scheme compared to the state-of-the-arts through two real-world case studies in social sensing applications.

II. Related Work

Social sensing has evolved into a vibrant area in sensor network research with widespread use of digital sensors, the ubiquitous Internet connectivity, and the advent of online social media. A comprehensive overview of social sensing applications is given in [14]. The truth discovery problem in social sensing is a key challenge that must be resolved for these applications to be useful [15]. Prior works saw significant progress in addressing this problem by making an important assumption: the source-claim relational dependency can be represented by some simplified functions [16], [17]. In contrast, this paper removes such assumption by developing a new multi-layer neural network approach that automatically learns the relationship between the reliability of sources and claim truthfulness from the observed data.

Neural networks are widely used for image recognition, and classification because images can be intuitively represented as multi-dimensional input vectors. In the machine learning and data mining communities, a large amount of work currently exist that relates to the topics of *fact-finding* and the majority of these approaches jointly compute the reliability of sources and claim truthfulness iteratively [18]. *TruthFinder* [8] established an unsupervised fact-finder to determine the reliability using a providers-fact network. *Hubs and Authorities* [19] used a simple fact-finding model that was based on strong linear computations for calculating what is known as reliability scores for sources and the claims they reported. There were also other fact-finding techniques that were used over time to enhance the basic frameworks mentioned above by analyzing dependencies or properties within the sources and claims [20].

III. Problem Statement

In this section, we define the truth discovery problem in social sensing. In particular, we consider a social sensing application where a set of M sources $S = (S_1, S_2, ..., S_M)$ collectively report a group of N claims $C = C_1, C_2, ..., C_N$. In this paper, we focus on binary claims (i.e., either true or false). Specifically, let S_i represent the i^{th} source and C_j represent the j^{th} claim. $C_j = \text{True}$ and $C_j = \text{False}$ represents the claim is true or false respectively. We define a *Sensing Matrix* SC, where $S_iC_j = 1$ when source S_i makes the claim C_j and $S_iC_j = 0$ otherwise.

The truth discovery problem in social sensing is to estimate the truthfulness of claims and the reliability of sources without prior knowledge of either of them. Let re_i represent the probability of source S_i to make true claims. Formally, re_i is defined as:

$$re_i = P(C_j = \text{True} | S_iC_j = 1) \qquad (1)$$

The goal of the truth discovery problem is to accurately estimate $Pr(C_j = \text{True}|SC)$ for all claims and re_i for all sources.

IV. Neural Network Solution

A. A Neural Network approach

We develop a neural network with the Dropout method to solve the above truth discovery problem. The main idea of the proposed neural network scheme is shown in Figure 1. The input to the scheme is a vector $S \in \mathbb{R}^M$ representing the sources that reported a claim. The output is a vector $O \in \mathbb{R}^2$ where O_0 indicates the claim is true and O_1 indicates the claim is false. To estimate the correctness of a claim, the network will learn a set of weight matrices $W_1, W_2, ..., W_z$ where z represents the amount of layers in the neural network architecture (e.g., the input layer is W_1 and the output layer is W_z). As the network gets deeper (i.e., more layers are added to the structure), more weight matrices W will be learned to model the *complex source-claim relational dependency*, the key challenge addressed by this paper.

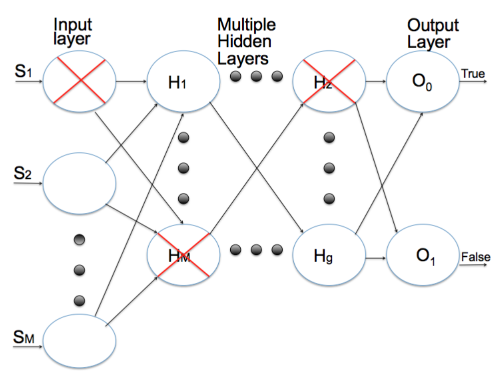


Figure 1. Overview of the Neural Network Scheme With Dropout. The nodes that are crossed out represent the ones that are excluded from the network based on Dropout to avoid over-fitting problem.

A critical step of the multi-layer neural network scheme is to train the weights W for each arbitrary layer in the network.

In particular, we represent the weight W as a matrix $W \in \mathbb{R}^{kl}$ where k is the dimensionality of the input to a specific layer and l is the dimensionality of the output to that layer. By using matrices, we can compute the product of the input vector S and hidden layers to obtain the output layer weights W_z. Once the weights are calculated, the correctness of a claim can be decided based on the results of O. The reliability of a source can then be computed as the ratio of correct to total number of claims made by the source.

All nodes on the input layer will be connected to the hidden layers which contain the weight matrix W to better learn the relationship between source reliability and claim correctness. The hidden layers will contain a vector b to adjust our weights to better match the desired source-claim relational dependency and allow faster convergence. Hidden layers are of smaller sizes than the input layer, which allows the model to learn different polynomials of the target function and forces the layer to learn a compact representation of the weights from the previous layer. Going from multiple hidden layers to an output layer, there is another matrix weight of dimensions $M \times 2$. Using the weights of the output layer, we can compute a score for each claim:

$$score_j = {v_{wj}}^T \cdot b \tag{2}$$

where v_{wj} represents the j column of the output layer W and b represents the bias. These scores can be further transformed into probabilities as follows.

$$output = \frac{exp(score_j)}{\sum_{i=1}^{M} exp(score_i)} \tag{3}$$

In this paper, we designed four layers to improve the truth discovery performance of the neural network approach. For the hidden layers, a rectified linear unit was added to better learn the complex relationship and non-linearities between source reliability and claim correctness.

$$f(x) = \begin{cases} x & : x > 0 \\ 0 & : x <= 0 \end{cases} \tag{4}$$

$$e^x/(e^x + 1) = \frac{1}{(1 + e^-x)} \tag{5}$$

The activation function for a rectified linear unit (RELU) is shown by Equation (4). We choose the rectified linear unit because it helps clip empirically set learning rates during training and it provides a simple derivative.We also show the derivative of the equation involved for smoothing the RELU which is known as the soft-plus function defined in equation (5). The output layer provides probabilities that can be used to decide whether a claim is true or false. The results of the output layer will be used to compute $p(claim_{correct}|claim_{guessed})$, which is the probability of obtaining the correct claim class ($claim_{correct}$) given the class guessed by the neural network ($claim_{guessed}$). The loss function can be seen as the cross entropy measurement as follows.

$$Loss = -log(p(claim_{correct}|claim_{guessed})) \tag{6}$$

Error is back-propagated through the network if the wrong prediction is made. The updated weights are computed via steepest descent minimization method during back-propagation using the current weights, the derivatives, and a variable η given as follows.

$$W_{ij} - \eta \times \frac{\partial loss}{\partial score_j} \times h_i = output - unit_j \tag{7}$$

where $output$ represents the studied claim and $unit_j$ is 1 if the actual claim prediction is correct and 0 otherwise.

The feed forward neural network uses tensor flow's stochastic gradient descent function to tune the weights for each layer in the network with multiple learning rates tested in the network to ensure the best results. The detailed learning rate and parameter tuning process is discussed in the next section. Once the neural weights are trained, the network will be able to assess the correctness of a claim from a set of users that reported that claim. Once we have claim correctness, we can calculate source reliability accordingly based on Equation (1).

V. Evaluation

In this section, we evaluated our neural network approach using two real-world social sensing datasets collected from Twitter data feeds. We chose Twitter as our social sensing platform because numerous unvetted social sensors (i.e, Twitter users) report their observations about the physical world in real-time [14] and the source-claim relational dependency is complex and unknown. In our evaluation, we compare the *Neural Network* scheme to a few representative baselines from recent literature. The first baseline we use is the *Regular EM*, which establishes a basic MLE model to solve the truth discovery problem under an assumption of a binomial distribution on the source-claim relational dependency [10]. The second baseline is *Conflict EM*, which extends the Regular EM scheme by considering the conflicting semantics of claims [21]. The third baseline is *Confidence EM* which extends the Regular EM scheme by considering the confidence of the sources in their claims [6]. The fourth baseline is the *TruthFinder*, which estimates source reliability and claim truthfulness based on a pseudo probabilistic model [8]. The fifth baseline is *Sums*, which assumes the source-claim relational dependency is simply linear [19]. The sixth baseline is the *Average_Log*, which explicitly considers both the credibility of sources and the claim count of the source in their methods [12]. The next baseline we use in our evaluation is *Voting*, which computes claim truthfulness by simply keeping count of how many times a tweet is reported through Twitter [22]. We also add a baseline called *Random* which shows the performance of a random guess of which claims are true or false. For the baselines that have parameters in their models (e.g., EM based schemes), we tune the parameters to achieve the best performance of those baselines for a fair comparison.

A. Data Collection

We implemented the neural network algorithm and prior baselines in the Apollo system [16], which is a social sensing framework developed specifically for gathering tweets and keeping track real-time events [23]. Some examples of these types of events include terrorists attack, tornadoes, political uprisings, earthquakes, riots, etc.

In our evaluation, we crawled two real world social sensing data traces from Twitter feeds about recent events. The first data trace was collected during the *Brussels Bombing* event which occurred on March 22, 2016, which was a series of terrorist attacks that left 35 dead in Brussels, Belgium. It is one of the worst terrorist attacks that happened in Europe this year. The second trace was collected in the midst of the *Baltimore Riots* event that occurred on April 14, 2015. It was a sequence of riots following the untimely death of an African American male in police custody in the city of Baltimore.

B. Data Processing

In our evaluation, we randomly sampled 2864 and 3029 tweets from the Brussels Bombing and Baltimore Riots data trace, respectively, due to limited amount of human powers to label the tweets. We asked five independent graders to manually labeled all claims using the rubric below and use majority voting of their votes to generate ground truth of claims.

- *True claims:* Claims from an event that can generally be observed by multiple independent human sensors and corroborated by credible sources external to online social media (i.e., the mainstream media).
- *Unconfirmed claims:* Claims that cannot be considered as true based on the requirements of true claims.

In our study, we observe that the unconfirmed claims may have some claims that are true but cannot not be independently verified by credible sources. Therefore, our method gives *pessimistic* performance bounds by labeling all unconfirmed claims as false. In order to evaluate the performance of our methods using real-world events, we performed the data pre-processing steps below:

Clustering: we use a variant of the K-means clustering algorithm and the Jaccard distance metric for micro-blog data clustering to group similar tweets into the same cluster [24]. In particular, the Jaccard distance is defined as $1 - \frac{X \cap Y}{X \cup Y}$, where X and Y define the group of words that appear in two compared tweets respectively. Therefore, the more similar the words the tweets share, the smaller Jaccard distance they have. In particular, we tokenize each tweet into a set of words and removed stop words as those can interfere with the Jaccard distance in the clustering step. We treated tweets that were clustered in the same cluster as a claim since they generally discuss the same statement.

Source-Claim Matrix Generation: we generate the SC Matrix by linking the sources with the claims that they make. We set the element S_iC_j in SC matrix to 1 if source S_i contributes a tweet that belongs to claim (cluster) C_j and 0 otherwise.

C. Truth Discovery Performance of Neural Network Approach

In this subsection, we evaluate the truth discovery performance of our neural network algorithm and compare it against the baselines methods discussed earlier in this section. In particular, we measure the performance of all compared algorithms by evaluating how accurately they can classify the true claims from the false ones using Precision, Recall, F1-measure, and Accuracy. The true positives and true negatives are the claims that are correctly classified as true or false respectively. The false positives and false negatives are the false and true claims that are incorrectly classified respectively.

We tune the parameters of our neural network model using a grid search approach to learn the best learning rate, batch and epoch size. The learning rate is the amount that matrix weights are updated during training, the batch size is the number of examples the network is shown before an update, and epoch size refers to the number of rounds that is used to update the network for the entire training dataset. In our experiments, we tried different values for the above parameters and found a batch size of 20, epoch size of 150 and learning rate of 0.025 gave us the best performance.

Figure 2 shows the results of the Brussels Bombing event trace. We observe that the neural network scheme continues to outperform the baselines. In particular, the performance gain of multi-layer neural network scheme compared to the best performed baseline is significant: *11%* in accuracy, *9%* in precision, *13%* in recall, and *14%* in F1-measure. We then repeated our experiments on the Baltimore Riots trace. The results are shown in Figure 3. Similar performance gains of the neural network scheme are observed.

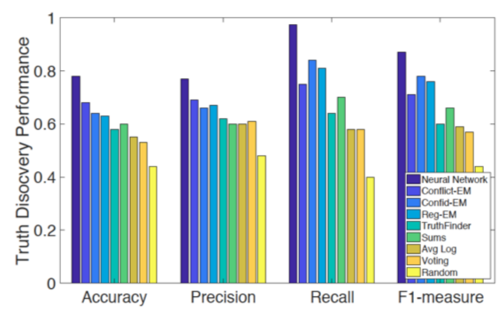


Figure 2. Brussels Bombing Trace

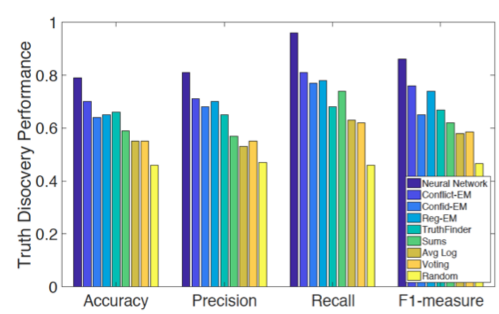
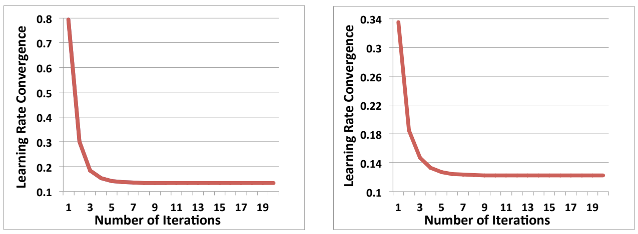


Figure 3. Baltimore Riots Trace

Finally, we study the convergence of the neural network scheme. In particular, we studied the learning rate changes with respect to the number of iterations (i.e. the loss at each step). The results are presented in Figure 4. We observe the neural network method takes only a few iterations to converge on all data traces. The above evaluation results demonstrate the effectiveness of using the proposed neural network approach to solve the truth discovery problem in social sensing applications.



(a) Brussels Bombing Trace (b) Baltimore Riots Trace

Figure 4. Learning Rate Convergence Analysis of Multi-Layer Neural Network Scheme with Dropout

VI. Conclusion

Our paper develops a novel neural network approach to solve the problem of truth discovery in social sensing applications. The proposed approach uses a multi-layer neural network to learn weights of layers from the sensing matrix and outputs correct estimates on the claim truthfulness and the reliability of sources. We evaluated the neural network scheme through four real world case studies using data collected from Twitter. The results presented in this paper are critical to this area of research as they lay out a concrete foundation for exploring the neural network based approaches to address the truth discovery challenge in social sensing.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. CBET-1637251, CNS-1566465 and IIS-1447795 and Army Research Office under Grant W911NF-16-1-0388. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

[1] Charu C Aggarwal and Tarek Abdelzaher. Social sensing. In *Managing and Mining Sensor Data*, pages 237–297. Springer, 2013.

[2] Jie Yin, Andrew Lampert, Mark Cameron, Bella Robinson, and Robert Power. Using social media to enhance emergency situation awareness. *IEEE Intelligent Systems*, 6:52–59, 2012.

[3] Francesco Calabrese, Laura Ferrari, and Vincent Blondel. Urban sensing using mobile phone network data: A survey of research. *ACM Computing Surveys*, 47, November 2014.

[4] Keiji Yanai. Urban sensing using mobile phone network data: A survey of research. In *ACM International Conference on Multimedia Retrieval*, ICMR ’02, page 65. ACM, 2015.

[5] Lance Kaplan, Murat Scensoy, and Geeth de Mel. Trust estimation and fusion of uncertain information by exploiting consistency. In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–8. IEEE, 2014.

[6] Dong Wang and Chao Huang. Confidence-aware truth estimation in social sensing applications. In *Sensing, Communication, and Networking (SECON), 2015 12th Annual IEEE International Conference on*, pages 336–344. IEEE, 2015.

[7] Zhou Zhao, James Cheng, and Wilfred Ng. Truth discovery in data streams: A single-pass probabilistic approach. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 1589–1598. ACM, 2014.

[8] Xiaoxin Yin, Jiawei Han, and Philip S. Yu. Truth discovery with multiple conflicting information providers on the web. *IEEE Trans. on Knowl. and Data Eng.*, 20:796–808, June 2008.

[9] Jermaine Marshall and Dong Wang. Mood-sensitive truth discovery for reliable recommendation systems in social sensing. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 167–174. ACM, 2016.

[10] Dong Wang, Lance Kaplan, Hieu Le, and Tarek Abdelzaher. On truth discovery in social sensing: A maximum likelihood estimation approach. In *The 11th ACM/IEEE Conference on Information Processing in Sensor Networks (IPSN 12)*, April 2012.

[11] Chao Huang and Dong Wang. Topic-aware social sensing with arbitrary source dependency graphs. In *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 1–12. IEEE, 2016.

[12] J. Pasternack and D. Roth. Generalized fact-finding (poster paper). In *World Wide Web Conference (WWW’11)*, 2011.

[13] Xiaoxin Yin and Wenzhao Tan. Semi-supervised truth discovery. In *WWW*, New York, NY, USA, 2011. ACM.

[14] Dong Wang, Tarek Abdelzaher, and Lance Kaplan. *Social Sensing: Building Reliable Systems on Unreliable Data*. Morgan Kaufmann, 2015.

[15] Daniel Yue Zhang, Chao Zheng, Dong Wang, Doug Thain, Xin Mu, Greg Madey, and Chao Huang. Towards scalable and dynamic social sensing using a distributed computing framework. In *Distributed Computing Systems (ICDCS), 2017 IEEE 37th International Conference on*, pages 966–976. IEEE, 2017.

[16] Dong Wang, Md Tanvir Amin, Shen Li, Tarek Abdelzaher, Lance Kaplan, Siyu Gu, Chenji Pan, Hengchang Liu, Charu C Aggarwal, Raghu Ganti, et al. Using humans as sensors: an estimation-theoretic perspective. In *Proceedings of the 13th international symposium on Information processing in sensor networks*, pages 35–46. IEEE Press, 2014.

[17] Daniel Yue Zhang, Rungang Han, Dong Wang, and Chao Huang. On robust truth discovery in sparse social media sensing. In *Big Data (Big Data), 2016 IEEE International Conference on*, pages 1076–1081. IEEE, 2016.

[18] Manish Gupta and Jiawei Han. Heterogeneous network-based trust analysis: a survey. *ACM SIGKDD Explorations Newsletter*, 13(1):54–71, 2011.

[19] J M Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.

[20] Dong Wang, Tarek Abdelzaher, Hossein Ahmadi, Jeff Pasternack, Dan Roth, Manish Gupta, Jiawei Han, Omid Fatemieh, and Hieu Le. On bayesian interpretation of fact-finding in information networks. In *14th International Conference on Information Fusion (Fusion 2011)*, 2011.

[21] Dong Wang, Lance M. Kaplan, and Tarek F. Abdelzaher. Maximum likelihood analysis of conflicting observations in social sensing. *TOSN*, 10:30, 2014.

[22] J. Pasternack and D. Roth. Knowing what to believe (when you already know something). In *International Conference on Computational Linguistics (COLING)*, 2010.

[23] Apollo-Toward Fact-finding for Social Sensing. http://apollo.cse.nd.edu/.

[24] Kevin Dela Rosa, Rushin Shah, Bo Lin, Anatole Gershman, and Robert Frederking. Topical clustering of tweets. *Proceedings of the ACM SIGIR: SWSM*, 2011.